Digital Preservation as an Albatross

James Currall, Peter McKinney and Claire Johnson
espida Project, University of Glasgow, Glasgow, Scotland

Abstract

'Digital Preservation' as a concept is an albatross. The complex and somewhat arcane nature of the practice has kept it from being embraced by those that perhaps need it most. Changes in terminology, misunderstandings of meanings and a lack of direct business planning have brought about a state of affairs that has the digital preservation community fighting the problem of technological obsolescence without sustained support from organisations that supposedly need it most.

Organisations care about ensuring their continued existence and profitability. Investment is only undertaken after reflection on business cases. In creating a business case most people focus primarily on cost, but there must be a counter-veiling focus on value. There is no point in making an investment unless it has worth to the investor. A good business case will display a strong understanding of the value of information objects that organisations create. Information professionals must ensure that their desire to ensure longevity of information is tied coherently and explicitly to that of the organisation's future and detail why the digital materials are of value to it. Exploring value in this way allows engagement with senior management as it wraps the need for action in the terminology of their strategic vision and allows for a strong and successful business case to be made.

Introduction

"The land of ice and of fearful sounds, where no living thing was to be seen. Till a great sea-bird, called the Albatross, came through the snow-fog and was received with great joy and hospitality. And lo! the Albatross proveth a bird of good omen, and followeth the ship as it returned northward through fog and floating ice."[1]

Digital Preservation - what is it? The preservation of digits? Is this the real task? As information professionals should we not centre our interest on the preservation of information rather than developing digital preservation as mysterious? Managers of enterprises are interested in managing the business - its processes, people, information and outputs. Digital preservation is not an end in itself, it is just a set of technical - ***how*** - things. In the early stages of any development, people tend to concentrate on the mechanics, rather than on that which is really important - ***what*** - and the purpose - ***why***. The result is that this activity has tended to be funded through short-term projects focusing on ***how***. What is now needed is for these practices to be seen as part of mainstream activity; which they surely are and to get their funding as part of the mechanisms that fund the other aspects of normal business activity. As yet few good business cases have been made for digital preservation *per se*. This is principally because most project proposals have concentrated on threats and ensuing costs, not ***why*** digital assets are to be preserved. As soon we start to consider ***why***, then we are talking about the value or benefit that such an investment might bring. Value is not absolute in that it is critically dependent on the question "***to whom?***" and "***when?***". In the case of an enterprise, this must be through a dynamic alignment with the purpose and strategy of the organisation.

It is our contention that the term '***digital preservation***' encourages a ghetto of practice and practitioner that hinders the embedding of overall good practice in information object creation and management into organisations - it is seen as an arcane art, something distinct, rather than something which should be a part of everyday activity. If we believe that information assets are in need of management and preservation, which we do, then the value of such activities in whatever medium the information is represented needs to be clearly expressed for decision makers, so that they understand what the magnitude (and other parameters of their investment) is, the value or benefit that derives from it and the risks inherent in making or not making the investment. This is the stuff of which business cases have been made traditionally for a wide variety of projects in government and commerce, but which presents considerable challenges when the benefits are not financial or tangible, but concern such difficult ideas as 'reputation', 'innovativeness', 'motivation' 'process capability' 'customer satisfaction' and 'cultural heritage' as they now do in information management in the digital age.

Business Case

Business cases are a concise non-threatening way of presenting investment opportunities. In order to decide whether or not to invest, the decision maker needs three pieces of information:

- the quantity of resource that must be invested (costs),
- what the investment is expected to deliver (benefits),
- the uncertainty in the size of the costs and whether or not the benefits will accrue (risks).

The decision to invest will only be made if these three factors fall within certain parameters (whether explicitly outlined or not). What is needed in information management is a simple (but not simplistic) model of cost, benefit and risk, that makes it straightforward to construct business cases that appeal to decision-makers in terms with which they and their stakeholders are familiar, rather than in the obscure language of 'digital preservation'. The ***espida*** project is developing just such a model and is testing it out in the University of Glasgow and elsewhere. Too many factors, not least the human factor, make producing the 'perfect business case' an impossibility, this model therefore does not attempt to do so, nor does it offer information professionals a specific tool to gain an advantage in the decision-making process - the same model could be used by anyone wishing to commend a course of action to decision-makers). Rather, the model presents a way of communicating effectively intangible value so that it can be understood and can have influence in investment decisions alongside harder, financial benefits. It is relatively simple to show how the purchase of a new machine will increase productivity and

therefore add value, it is much more difficult, as the dotcom bubble demonstrated, to do the same for information objects.

Cost

The costs of 'doing something' can be determined relatively easily once certain assumptions have been made. Our approach to cost is therefore uncomplicated and does not seek to replicate the excellent work that has already been done in many places to produce tools that estimate the costs and economies of scale of specific information preservation actions. The model points to tools that can help pull together the financial costs but does not offer direct guidance on the use of these. Costs will include the direct costs of the equipment, software and staff and these will consist of both Capital costs and Revenue (or recurrent) costs. Obviously these costs must be modeled over the length of time that the information must be preserved and when appropriate finally destroyed. They should not cover just the initial period of establishment of a repository or the ingest of material.

In our work, we are not prescriptive about the basis of costings, but we do require that the basis on which figures are calculated or given is made explicit, so that decision-makers can judge the reliability and whether or not they are realistic.

Value

Benefits are the primary reason for making an investment, financial or otherwise. It is of paramount importance to be able to communicate effectively the benefits gains to those being asked to make an investment and their stakeholders. It is clear that in talking about information resources we are, for the most part, talking about intangible assets and that value based simply on financial measures is inadequate. Whilst it is possible to try to convert all value to a monetary value, doing so can be misleading, time-consuming, unrepresentative and counter-productive. This was the painful lesson of the dotcom boom. How best then to communicate the benefits and thereby attract investment?

The latent value that is capable of producing benefit is multifaceted. Intangible value is therefore very difficult to deconstruct and communicate. Exploring the world of intangible value, led us to a technique used for business performance (amongst other things) called the Balanced ScoreCard (BSC), developed by Kaplan and Norton at the Harvard Business School.[2] We have repurposed Kaplan and Norton's concept in a fairly cavalier way, but according to a range of external assessors of the project the result seems to be promising.

A basic premise is that value is not absolute, but must be seen in relation to the strategic goals over time of the organisation bearing the costs. The Balanced ScoreCard was designed as a performance measurement tool to allow senior management to monitor the performance of their company (or parts of it) in line with its strategic aims. The adaptation has two roles: it allows producers of a business case to consider and communicate all areas of value from which the organisation will reap benefit; and, it also acts as an investment measurement tool for senior management.

The ***espida*** Balanced Score Card (*sic*) is illustrated below and views value from four distinct perspectives, of which only one is concerned with financial value. It explores the perspectives of :

- ***the customers*** (what value does the information asset bring for customers?);
- ***the internal business processes*** (does the information asset help the way the organisation carries out its business?);
- ***the innovation and learning*** perspective (can the information asset help the organisation develop?); and finally
- ***the financial*** perspective (does the information asset bring financial savings or income?).

These perspectives can be broken down into constituent elements (see Appendix 1). It is vital that these perspectives and elements are fully aligned with the strategic aims of the funding organisation: the business case as a whole (be it buying a digital repository, resources for cataloguing or records management) must show value that will be of benefit to the organisation otherwise the information asset is not relevant to that organisation and there will be no investment. The elements listed in the appendix reflect the directions from which value in information assets comes in the case of the University of Glasgow, but whilst some of these are appropriate to a wide range of activities and organisations, others are quite specific to a research-led university. This alignment with strategic objectives (broadly specified) is why 'digital preservation' is an albatross. It has not placed sufficient emphasis on the question '***why***', it has presumed that the materials at risk are of concern to the organisation without specifically asking or answering the question '***why should these information assets be preserved in the first place?***'. The focus has been on funding an action, rather than the benefit of that action.

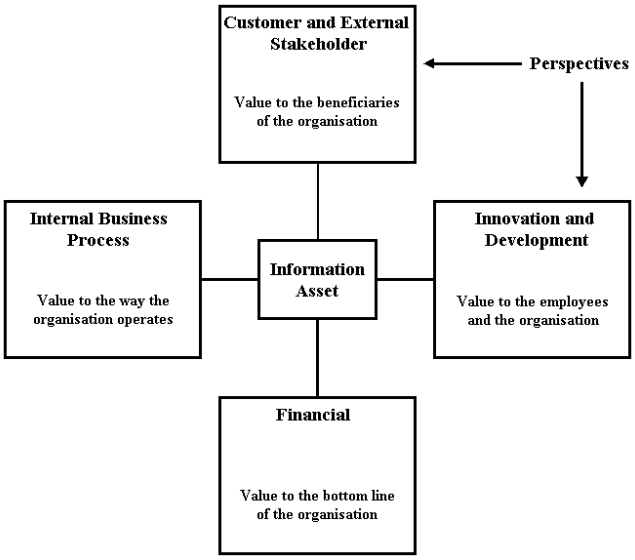

Figure 1. *espida Balanced Score Card (BSC)*

Clearly in this formulation, an information asset might produce negative value if it reduces rather than increases one of the Score Card elements. It is this capacity to produce negative value that records managers sometimes refer to as risk. In the ***espida*** model value can be negative or positive, but investments will be targeted at courses of action that produce a net positive value rather than a net negative value. We can observe much the same process at work with physical assets, where the pay back from the cost of an investment is less than the outlay.

Value Metrics

With regard to metrics, we have rejected the idea of having globally applicable measures of value, since agreeing on common currencies would be an impossible task and, in any event, will change over time. Instead we put the onus on the 'project' proposer to indicate how they will know that the value indicated has been realised. In indicating suggested 'metrics' to the funder, the proposer has to strike a balance between talking the 'project' up and setting targets that cannot be delivered. Our evaluators have been very positive about this approach that again has its equivalent in tangible assets, for example the output of a specific process.

Proposing a metric with which to measure the attainment of the benefit allows monitoring of the investment, acts as an important way of communicating of the strength of value and gives the proposer buy-in. As an example, if a business case was being made for the construction of a trusted digital repository. It might be that this would be deemed to bring strong value in relation to information accessibility and the measure of success might be attracting 10,000 hits per month. This measurement offers senior management a scale to the benefit that the proposer thinks is attainable for the investment and allows them to monitor the success of the investment subsequently. This measurement framework helps the proponents to keep a sense of perspective.

Uncertainty

Uncertainty or risk plays a major role in any investment decision. Reaction to uncertainty is dependant on the risk appetite of decision maker. Those that are risk adverse will view a medium risk in a very different way than those that are risk takers (it can often be seen that the public sector is more risk adverse than the private sector). It is very important therefore to not only be able to communicate what the degree of risk is, but also what the perception of these risks is.

The term Risk Analysis is often employed to mean - producing a horrifying list of possible contingencies that we must try to avoid. Any formulation of risk means very little if the following three components are not all present for all the items on such a list:-

1. a contingency (the element itself in this case)
2. the likelihood (of a element being realised)
3. the impact (the size of the element in terms of what it delivers to the organisation)

Possible courses of action to mitigate the risk can be formulated to remove the contingency, reduce the likelihood or lower the impact.

Preserving information assets or disposing of them involves two sort of uncertainty:

- the proposed benefits failing to materialise in spite of the investment,
- possibility of making the wrong decision.

The risk is in the first case is that for one of a number of reasons the information assets were over-valued and in the second case it is concerned with the consequences of the decision itself being wrong in relation to the likelihood of this being so.

In our model we accept that no investment is certain to produce the benefit slated for it and ask the propser to indicate likelihood. We recognise that if large benefits are claimed for an investment, then the risk of the investment delivering will be lower than if the benefits suggested are lower and would therefore counsel realism over hyperbolic optimism when assessing the benefits of a project, an observable tendency in our experience in the digital order.

The uncertainty within the model is concerned only with the risks that impact on the decision making process, this means that risks that come into play during the actual investment are not considered in detail. (These should be taken care of by budgeted contingencies.) Primary uncertainties in the model are therefore whether the outlined values will come to pass and if the recurring costs outlined are likely to rise or fall.

Making Decisions

In our experience decision-makers need to be able to:

- recognise the different types of value (1 above),
- have a realistic idea of the likelihood of any element being realised (positive or negative) (2 above), and
- have some measure of the magnitude of the value (3 above), in order to judge whether or not the investment is sound.

The judgment itself clearly depends on what the competing demands for resources are and how risk averse the organisation concerned is.

The current state of our methodology is that we have templates for Value and Cost, that have been tested by a number of external evaluators and refined to take account of their feedback. The Value template that we are using in the University of Glasgow contains the perspectives and elements in appendix 1, aligned directly with the strategic goals of that Institution with the help of the primary architects of the University Strategic Plan. Within this template, the strength of the value (be it positive or negative), the likelihood of the organisation receiving benefit from the investment and the timescale of this benefit coming to pass are all required. Crucially, it also requires measurement metrics to indicate what the benefit will be. (The model includes mechanisms for modeling value over time as its shape can change dramatically and that can have important repercussions for investment decisions.)

These templates are used by the proposer of a 'project' or 'activity' to indicate what costs they expect and to provide a profile of the expected returns in value (positive and negative) terms. The templates concentrate attention on:-

- the nature of benefits and dis-benefits that might be expected,
- the size of these returns - in relative terms within the project not global terms, on the grounds that this allows the 'funder' to see clearly where the main elements of value are,
- the timescale of returns,
- getting the proposer to indicate how the value can be measured, which provides a basis for both deciding whether or not to invest and for evaluating the investment in a post-implementation review at later stages,
- making the proposer 'show their working' and explain how they come to their conclusion that the project is worthwhile.

The reality of decision-making is that those making the decisions will have to choose between a variety of very different projects when deciding how to invest the money available. The cost and benefit templates make it much easier for those involved to see clearly what they might get for their money and to ask for clarification where it is needed.

Conclusion

Whilst we set out to provide a way of indicating value of digital resources, it has since become apparent that there are a large number of areas where the developing methodology might be appropriate:-

- ***business cases*** - for projects involving intangible benefits,
- ***impact of change*** - for comparing a change proposal with the current situation,
- ***evaluation of proposals*** - providing a 'level playing field' for comparing a range of different proposals, where benefits are often expressed by the proponents in hyperbolic prose, making comparison and evaluation difficult.

In addition discussions with a number of people with expertise in different areas indicate that there is clear interest in the applicability of this type of approach in the following areas:-

- Digital Preservation
- Information Technology/Information Systems projects
- Library and Archival Collections
- Cultural Heritage

'Should I shoot the albatross?' the mariner could have asked himself (and possibly should have). What may happen if I do not shoot it and what may happen if I do? Whether the mariner carried out what would now be known as option appraisal is uncertain, but within the context of the model the options are the same: to do or not to do something.

What we are arguing is that for the preservation of information in digital form to become part of mainstream activity, it needs to be subject to the same sort of justification as any other form of investment. So long as it is seen as a separate activity, carried out by a group of 'anoraks' who 'specialise in that sort of thing', it will remain under-funded, under-valued and mis-understood by the people who run organisations. We propose a practical way of making this a reality. Of course in advocating a course of action that might be seen as "shooting the albatross", we risk the fate of the Ancient Mariner:

"And ever and anon throughout his future life an agony constraineth him to travel from land to land and to teach by his own example, love and reverence to all things that God hath made and loveth."[1]

References

[1] Coleridge, S.T. *The Rime of the Ancient Mariner*. (Dover Publications Inc. 1970)

[2] Kaplan, R.S., and D.P. Norton. "The balanced scorecard - measures that drive performance". (*Harvard Business Review* 1992) 70, 58-63.

Author Biography

James Currall is the Director of the espida project, he is a statistician, whose current job involves the planning and development of IT services, he is also researching into information access and the nature of digital objects and the problems associated with their management and retention.

Peter McKinney is Research Officer for the espida project. Previously, he was Coordinator of ERPANET, a key international player in the area of digital preservation and was part of the team that developed the NINCH Guide to Good Practice in the Digital Representation & Management of Cultural Heritage Materials.

Claire Johnson leads the records management team at the University of Glasgow and is currently responsible for Freedom of Information across the Institution and is researching various aspects of records management including the impact of FoI on the public sector.

Appendix 1: Balanced Score Card Elements

Innovation and Development Perspective

- Intellectual capital
- Motivation, fulfillment and satisfaction (of staff)
- Quality and potential of research
- Quality and potential of teaching
- Responsiveness to change

Internal Business Process Perspective

- Information accessibility
- Efficiency of operation and productivity
- Effectiveness of decision making
- Process potential and organisational flexibility
- Compliance with legislation and regulation

Customer and External Stakeholder Perspective

- Contribution to culture and community
- Reputation, brand and customer confidence (in all who deal with the University and in the public at large)
- Customer satisfaction and service delivery (students, parents, public, etc.)
- Academic attractiveness (to potential students, staff, academic partners and funding agencies)
- Commercial attractiveness (to potential sponsors and collaborators)

Financial Perspective

- Income generation
    - selling assets
    - licensing/rights to assets
    - teaching and research
    - contracts, grants, fees, donations, etc.
- Cost saving
    - labour, time
    - space
    - direct expenditure